Optimization of Campaign Budget Allocation using

Genetic Algorithms

Shritej Shrikant Chavan, Shaurya Tripathi, Shashank Ravishankar, Shaiba Jhunjhunwala, Sarthak Vajpayee

[sxc220078@utdallas.edu](mailto:sxc220078@utdallas.edu), [sxt220033@utdallas.edu](mailto:sxt220033@utdallas.edu), [shashank.ravishankar@utdallas.edu](mailto:shashank.ravishankar@utdallas.edu), [shaiba.jhunjhunwala@utdallas.edu](mailto:shaiba.jhunjhunwala@utdallas.edu), [sxv220020@utdallas.edu](mailto:sxv220020@utdallas.edu)

Abstract

*Optimization of resource allocation for online marketing campaigns is crucial in the digitization era and a standard problem in Operations Research. In this research paper, given the historical data for marketing campaigns, we attempt to predict digital engagement such as click-through rate using a regression and optimization approach to optimally allocate the spending. To evaluate the optimization of non-linear, non-differential objective functions we employ genetic algorithms as a robust approach. Our algorithms iteratively find the optimal solution, we allocate most of the budget to social channels across Audience type3 and 2. Finally, we discuss some underlying limitations in our approach and propose some recommendations for future work.*

Introduction

Given the ubiquity of digital platforms, the domain of digital marketing has boomed and increase the scope of marketing channels. To create brand awareness, companies spends a fortune on marketing their brands, products and services across different channels where they cater to various audiences. Therefore, it’s essential for them to allocate their budge optimally so that they can maximise their cumulative return. In this article, we explain our approach to maximize the digital engagement of a campaign across channels and audiences while satisfying some business constraints. To solve the complex optimization formulation problem, we leverage genetic algorithms and find the optimal budget allocation as % of the total budget.

**Solutions**

Given the task of optimizing the budget of $1M for the campaign, we first chose to predict the metric that defines the campaign return and then formulate an optimization approach that tries to maximize that return and in turn optimally allocate the budget for the campaign across different channels and audience types.

**Regression:**

Considering the futility of the ads that have zero spend and zero impressions, we chose to exclude that from our training data. Moreover, given that we needed to optimize our spend across only two variables namely Channel and Audience, we only use them as our independent variables in our regressor on top of spend.

The metric we used to quantify our campaign return is Click-through rate. It is defined as Number of Clicks by Number of Impressions.

Then we build a regression model of the following form:

Note: Though this above regressor is written in Linear Regression form, we chose Random Forest Regressor to predict Impressions and XGB Regressor to predict click which was fitted on the training data due to poor performance of Linear Regression model.

**Optimization:**

To optimally allocate budget for the campaign across 3 channels and 5 audience types, we will have 15 decision as Sij where i – channels, j- audience type. Here, for simplicity we define spend variables as percentage of the total budget therefore their sum equates to 100.

After predicting the CTR as a function of spend, we will use that our objective function:

Objective function:

Constraints:

Given the non-linearity of the objective function, we couldn’t use simple linear optimization algorithms like Simplex. Moreover, since the objective function is non-differentiable as well, it’s inconsistent to use Lagragian Multiplier. Hence considering the complexity of the problem and possible non-convex solution space we explore the novel method of solving non-differentiable, non-linear objective using Genetic Algorithms that is based on the principles of natural selection and genetics.

**Optimization Implementation:**

This is an iterative method that starts by initializing set of population of 100 random individuals (possible solutions), where each individual in our case is n-bit binary representation of list of spend variables. Each individual is evaluated by using a fitness function and then ranked accordingly. In our fitness function is same the objective function defined above.

This algorithm is ran for 1000 generations where top-ranked parents for that generation (selection) are chosen to breed (crossover) using bitwise XOR and XNOR function and create children. Then given the mutation rate of the population, we mutate each children is random value in [0,1] < mutation rate (mutation). If each case we make sure the constraints are not violated. The equality constraint is check by calculating the sum of a solution and if it falls beyond the threshold (±10) we carry the process which might be either mutation or crossover. However, if it’s inside the threshold then we use repair function to slightly modified our child so that best possible solution is not completed ignored.

**Model evaluation metrics:**

1. **Regression:**

R-squared of Random Forest Regressor Impressions = 0.7343

R-squared of XGBRegressor Clicks = 0.6589

1. Optimization: Fitness Function = 8.2446

Results

The results are based on the optimization of the Click-through rate. The distribution of $1 million spending budget is as followed-

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
|  | Audience Type 1 | Audience Type 2 | Audience Type 3 | Audience Type 4 | Audience Type 5 |
| Programmatic Channel | 2% | 3% | 12% | 3% | 2% |
| Search Channel | 17% | 2% | 2% | 2% | 15% |
| Social Channel | 2% | 14% | 20% | 2% | 2% |

Justification of solution:

1. The algorithm allocated 2% spend allocation (minimum requirement) to Programmatic Display Channel (Audience Type 1), Search Channel (Audience Type 2, 3 & 4) and Social Channel (Audience Type 1, 4 & 5), since our dataset lacked data for these variables.
2. Social Channel, Audience Type 3 was allocated the highest spending budget of 20%. This makes sense as the variable had the highest average CTR of 0.63, while Programmatic Channel, Audience Type 5 was allocated the minimum spend of 2%, which had the lowest average CTR
3. The remaining variables have budget spend allocations close to their average CTRs, which completely satisfies our spending constraint of 100% and maximizes the Click-Through Rate

Conclusions

Genetic algorithms were used to allocate advertising budgets to various channels based on audience types and their average CTRs. The highest spending budget of 20% was allocated to Social, Audience 3 with the highest average CTR, while the lowest spend of 2% was allocated to Programmatic Channel, Audience 5 with the lowest CTR. The remaining variables were allocated budgets close to their average CTR to maximize overall CTR and satisfy the 100% spending constraint.

The limitations of our approach are as follows:

1. Metrics for Campaign return – Need for experimention our approach using more robust metric to quantify campaign return. Examples Conversion Rate, Pay-per-Click, Impressions or some weighted combinations of all the above.
2. Feature Engineering- Lack of feature engineering to create more meaningful features to better predict our regression model
3. Tune Hyper-parameters – Tune hyperparameters for regression models including XGBoost and Random Forest regressor as well as genetic algorithms.
4. Fitness Function – Define fitness function as weighted sum of some metric for campaign return rather than plain sum

References

Katoch, S., Chauhan, S.S. & Kumar, V. A review on the genetic algorithm: past, present, and future. <https://doi.org/10.1007/s11042-020-10139-6>

A new genetic algorithm approach for optimizing bidding strategy viewpoint of profit maximization of a generation company, <https://doi.org/10.1016/j.eswa.2011.05.015>

User Response Learning for Directly Optimizing Campaign Performance in Display Advertising. Kan Reny, Weinan Zhangy , Yifei Rong♯, Haifeng Zhang♮, Yong Yuy, Jun Wangz